Theory and simulation of macromolecular crowding effects on protein folding stability and kinetics

Jeetain Mittal[1,*] and Robert B. Best[2,†]

[1]*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA*

[2]*Department of Chemistry, University of Cambridge, Cambridge, UK*

(Dated: February 26, 2009)

Abstract

We investigate the effect of macromolecular crowding on protein folding, using purely repulsive crowding particles and a self-organizing polymer model of protein folding. We find that the thermodynamics of folding for typical α-, β- and α/β-proteins are well described by an adaptation of the scaled particle theory (SPT). In this approach, the native state, transition state, and the unfolded protein are treated as effective hard spheres with radii approximately independent of the size and concentration of the crowders. The same model predicts the effect of crowding on the folding barrier and therefore refolding rates with no adjustable parameters. A simple extension of the SPT model, assuming additivity, can also describe the behavior of mixtures of crowding particles.

In contrast with conventional laboratory experiments conducted under dilute conditions, protein folding in a cell occurs in a dense environment consisting of various other macromolecules, commonly referred to as "crowders" [1]. Although the detailed interactions of the protein with the crowders may be quite complex, the primary physical effect of macromolecular crowding on the protein folding reaction is to reduce the volume available to the protein by that occupied by the crowders [2, 3, 4, 5, 6, 7]. A complete theoretical understanding of the excluded volume effects will greatly enhance our ability to interpret experiments, as well as all-atom simulations, and to develop coarse-grained models [8], of concentrated protein solutions.

Several theories have been put forward to predict the effect of crowders, modeled as impenetrable repulsive particles, on the folding free energy. However, these theories provide strongly contrasting predictions for the quantitative and sometimes even qualitative effects of crowding. For example, by treating the folded and unfolded proteins as effective hard spheres, Minton utilized scaled particle theory (SPT) to estimate the change in folding free energy as the difference between the insertion free energy for the folded and the unfolded states. The SPT free energy of inserting a hard sphere of radius R in a hard-sphere fluid of particle radius $R_{\rm c}$ is [9], $\beta F = (3y + 3y^2 + y^3)\rho + (9y^2/2 + 3y^3)\rho^2 + 3y^3\rho^3 - \ln(1 - \phi_c)$, where $\beta = 1/k_{\rm B}T$, T is the temperature, $k_{\rm B}$ is the Boltzmann's constant, $y = R/R_{\rm c}$, $\rho = \phi_c/(1 - \phi_c)$, and ϕ_c is the fluid volume fraction. This theory predicts a rather strong effect of macromolecular crowding on the folding free energy [10], with a monotonic increase in stability with increasing crowder packing fraction ϕ_c.

An alternative theory proposed by Zhou also uses SPT for the the effect of crowders on the folded protein, but the free energy of the unfolded protein is calculated using an elegant model of the unfolded chain as a random walk in the presence of a spherical trap [11]. For a Gaussian chain of radius of gyration $R_{\rm g}$, with $z = R_{\rm g}/R_{\rm c}$, the change in free energy is $\beta F = 3\phi_c z^2 + 6\pi^{-1/2}\phi_c z - \ln(1-\phi_c)$. This model predicts a much weaker effect of crowding on the unfolded state free energy because the unfolded polypeptide can access the void space between the crowders. In addition, a maximum in stability is predicted as the crowder packing fraction is increased. The extension of this theory to binary mixtures of different size crowders leads to an intriguing conclusion that there will be an optimum mixing ratio of the two components to achieve maximum protein stability [12]. One may expect that different crowding theories will work better under certain conditions of crowder packing

TABLE I: Characteristics of the proteins considered.

Protein	$a_{\rm N}$ [Å]	$a_{\rm U}$ [Å]	$a_{\ddagger}$ [Å]	$R_{\rm g}^{\rm U}$ [Å]
prb	9.5	12.75	11.4	15.9
proteinG	10.6	16.0	-	20.0
TNfn3	13.0	21.5	-	28.6

fraction and size. However, the boundaries for the validity of these theories in parameter space are relatively unknown.

In this letter, we investigate the appropriate theoretical description of purely repulsive crowders, using molecular simulations of a coarse grained folding model in a bath of crowders. We consider two-state proteins from the three main protein structural classes: all-α (prb [13]), all-β (TNfn3 [14]) and α/β (protein G [15]), which are described using a self-organising polymer model (or "Gō-like" model) [16, 17]. In our model, each amino acid residue is represented by a single particle and a standard procedure is used to build the potential from the experimental native-state structure [18]. Interactions between the contacts present in the native state are treated as attractive and all others as repulsive, an approximation motivated by the funneled nature of the folding free energy landscape [19]. Previous studies have shown that this type of simplified model can indeed capture relevant features of protein folding [20], such as mechanism [21] and kinetics [22, 23]. The repulsive interactions between a crowder and the protein or other crowders are given by the pair potential $V(r) = \epsilon[\sigma_{\rm ref}/(r - \sigma + \sigma_{\rm ref})]^{12}$, where r is the distance between particle centers, $\epsilon = 1$ kcal/mol sets the energy scale, σ is the hard core overlap distance, and $\sigma_{\rm ref} = 6$ Å is a reference diameter: for $\sigma = \sigma_{\rm ref}$, $V(r)$ reduces to a more familiar form. We define σ between the pair (i, j) as, $\sigma = R_i + R_j$, where R_i, R_j are the radii of either crowder particles $R_{\rm c}$, or of the various protein residues R_p. We use Langevin dynamics simulations with a time step of 10 fs and a friction coefficient of 0.2 ps^{-1}, using the BBK integrator [24] in the CHARMM simulation package [25]. Cubic periodic boundary conditions with a primary cell size of 100 Å were employed. To speed up equilibration at a given temperature, we use replica exchange moves every 30 ps between 12 replicas which are each biased using an umbrella potential of the form $V_i(Q) = 0.5\kappa(Q - Q_i)^2$, where Q is the fraction of native contacts, $0 \leq Q_i \leq 1$ is the target Q value for replica i and $\kappa = 300$ kcal/mol is the force constant. The required

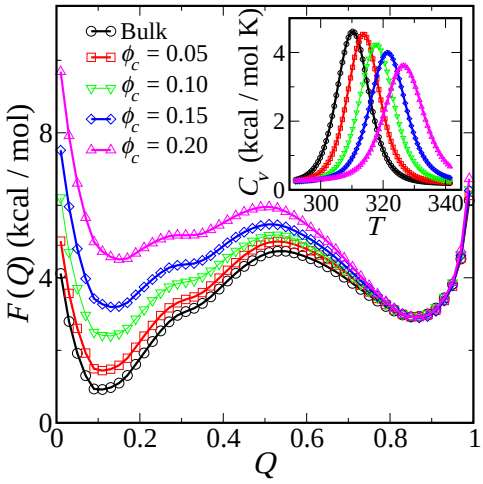

FIG. 1: Effect of crowding on folding free energy surface of protein G. The potential of mean force along the reaction coordinate Q, the fraction of native contacts, is shown for bulk conditions and with different crowder packing fractions ϕ_c for crowder size $R_c = 8$ Åat 320 K. The curves have been shifted by an arbitrary constant to match the free energy of the folded state. Inset: Heat capacity $C_v(T)$.

thermodynamic information is extracted from simulations with different umbrella potentials and temperatures using the weighted histogram method (WHAM) [26, 27].

For the model proteins considered here, Q is nearly an optimal coordinate for identifying transition states and also capturing the dynamics of protein folding with a diffusive Markovian model [23, 28, 29]. Here, we project folding in our systems onto Q, and use it to identify transition states as well as unfolded and folded free energy minima. To estimate the crowding-induced changes on the folding free energy surface, we construct the potential of mean force (PMF) as a function of Q, defined as $\beta F(Q) = -\ln[P(Q)/\Delta Q]$, where $P(Q)$ is the equilibrium probability of observing configurations between Q and $Q + \Delta Q$. Figure 1 shows representative free energy profiles in bulk ($\phi_c = 0$) and under crowded conditions ($\phi_c > 0$, $R_c = 8$) for protein G. The destabilization of the unfolded state due to the presence of crowders is clearly visible from an upward shift in the curves near the unfolded basin ($Q \approx 0.1$). For high crowding packing fractions ($\phi_c > 0.10$), there is also a slight destabilization of the transition state ($Q \approx 0.5$) with respect to the folded basin ($Q \approx 0.9$). Fig. 1 (inset) shows the heat capacity C_v curves for the same state points, as obtained from the WHAM analysis. The maximum in C_v is defined as the folding temperature T_f. The shift

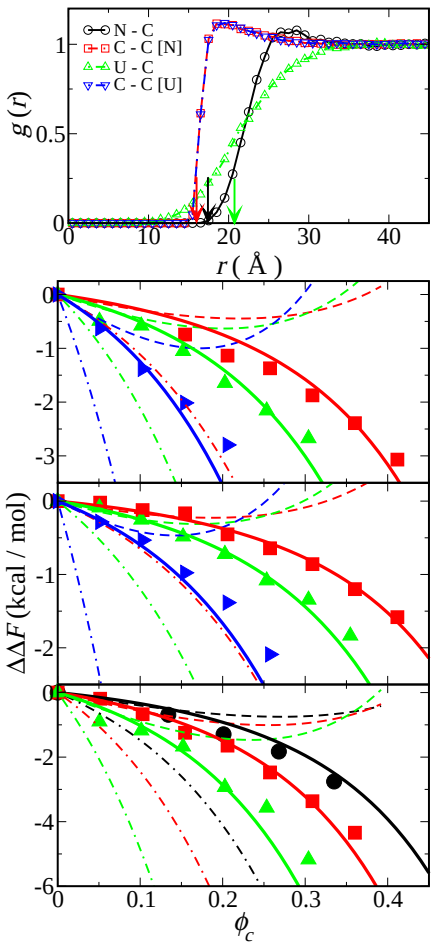

FIG. 2: Shift in protein stability with crowder size and packing fraction. Top panel: pair distribution functions $g(r)$ between crowders (C) and the native (N) and unfolded (U) protein (C-C [U] is the C-C $g(r)$ in the presence of U), for prb with $R_c = 8$ Å, $\phi_c = 0.05$. Arrows indicate corresponding sums of hard sphere radii. Lower panels: change in folding free energy with crowding: *upper*: protein G, *middle*: prb, *lower*: TNfn3. Symbols: simulations (errors smaller than symbol size) for $R_c = 20$ Å (circle), 16 Å (square), 12 Å (triangle up), and 8 Å (triangle right); dashed lines: predictions from Zhou theory [11]; dot-dash lines: predictions from Minton theory with $a_U = \sqrt{5/3} R_g^U$; solid line: result of using a smaller hard sphere for the unfolded state, motivated by the form of the U-C $g(r)$.

in T_f toward higher temperatures with increasing ϕ_c demonstrates the stabilizing effect of the crowders on the folded protein relative to unfolded.

Figure 2 shows the effect of crowding on the free energy of folding, $\Delta\Delta F = \Delta F_{U-N}(\phi_c = 0) - \Delta F_{U-N}(\phi_c)$, at $T = 320$ K for prb and protein G and $T = 300$ K for TNfn3, where

$\Delta F_{\mathrm{U-N}} = -k_{\mathrm{B}}T\ln(\int_0^{Q_\ddagger} \mathrm{e}^{-\beta F(Q)}dQ/\int_{Q_\ddagger}^1 \mathrm{e}^{-\beta F(Q)}dQ)$, is the difference between the unfolded F_{U} and native F_{N} free energy, $Q_\ddagger$ is the location of the transition state along Q. We find a monotonic increase in stability with the packing fraction ϕ_c for all proteins and crowder sizes. For a given ϕ_c, the smaller crowders are more strongly stabilizing because there are effectively fewer voids of the size of the protein. The prediction of Zhou's theory using the folded protein radius a_{N} and unfolded protein radius of gyration $R_{\mathrm{g}}^{\mathrm{U}}$ computed from the simulations (Table I) gives excellent agreement with the simulation data in the limit $\phi_c \to 0$ (Figure 2). For smaller crowders and higher packing fractions, however, the theory predicts too small a stabilization, most likely due to the neglect of excluded volume in the unfolded chain.

On the other hand, the Minton theory predicts a stabilization much greater than found in the simulations over the full range of ϕ_c. This theory treats the unfolded protein (with radius of gyration $R_{\mathrm{g}}^{\mathrm{U}}$) as an "equivalent hard sphere" of uniform density, i.e. with radius $\sqrt{5/3}R_{\mathrm{g}}^{\mathrm{U}}$. However, as indicated by the pair distribution function $g(r)$ between the unfolded protein and the crowders (Figure 2), the unfolded protein is quite soft, so that a smaller choice of hard sphere radius may be more appropriate. Remarkably, we are able to fit the data over the full range of ϕ_c and crowder size (which varies by a factor of 2) to the SPT model by using the radius of the unfolded state a_{U} as a single adjustable parameter. This radius is indicated on the plot of $g(r)$ in Figure 2. At high ϕ_c the smallest crowders are better able to penetrate the protein and the fit could be marginally improved by using a slightly smaller a_{U}. This treatment of the unfolded protein as a hard sphere is motivated by the highly successful Asakura-Oosawa theory [30] of polymer-colloid mixtures so that soft sphere interactions can be used between the polymer and crowder [31] which in turn can be mapped onto a hard-sphere system. In our case, we find that the effective hard-sphere radius a_{U} of the unfolded chain is $\approx 80\%$ of $R_{\mathrm{g}}^{\mathrm{U}}$.

We have also studied mixed macromolecular crowding with binary A:B and ternary A:B:C mixtures for the α-helical protein prb. For A:B mixtures, keeping the packing fraction of the type "A" crowders (radius 8 Å) fixed at $\phi_c^A = 0.05$, we varied the fraction of type "B" crowders (radius 12 Å). The simulation results for the change in folding free energy for different mixtures are shown in Figure 3. To test whether the effect of mixed crowding is simply additive and can be easily estimated from the pure crowder simulations, we also calculate the change in folding free energy from the following additivity *ansatz*,

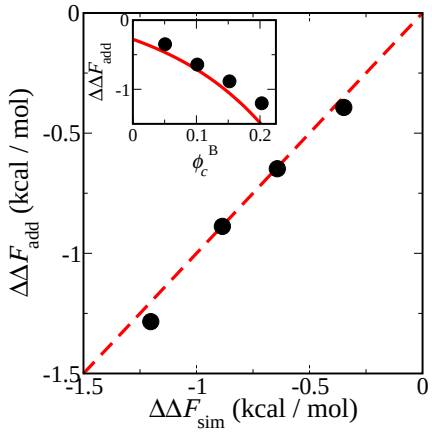

FIG. 3: Scaled particle theory prediction for binary mixtures of 8 Å and 12 Å crowders. With the packing fraction of particle type A fixed to $\phi_c^A = 0.05$ and the fraction of particle type B, $\phi_c^B = 0.05$ - 0.20, the change in folding free energy of prb from simulations $\Delta\Delta F_{\rm sim}$ is shown as a function of change in folding free energy $\Delta\Delta F_{\rm add}$ from the additivity *ansatz* (circle) as discussed in the text. The inset shows simulation data (circle) along with SPT model predictions (line) using the additive model.

$\Delta\Delta F_{\rm add}(\phi_c^1, \phi_c^2, ...\phi_c^N) = \sum_i x_i \Delta\Delta F_i(\sum_i \phi_c^i)$, where index i runs over N different types of crowding particles, $x_i = \phi_c^i / \sum_j \phi_c^j$ is the fraction of crowder type i in the mixture, and ϕ_c^i is the volume fraction of i. The predictions of this additive model $\Delta\Delta F_{\rm add}$ for A:B mixtures are in extremely good agreement with the simulation results $\Delta\Delta F_{\rm sim}$ as shown in Fig. 3. Moreover, for an A:B:C mixture of crowder radii 8, 12, 16 Å with volume fraction $\phi_c^i = 0.05$ of each component, the agreement between the simulation result (0.55 kcal/mol) and the additive model (0.54 kcal/mol) is remarkably good. Therefore, it should be possible to estimate the effect of mixed crowding due to several types of repulsive crowder by utilizing single crowder results. Applying additivity to our SPT model predictions for single crowders also provides reasonable estimates for $\Delta\Delta F$ (Figure 3 inset). Our results and the additive model for mixed crowding do not predict an optimal mixing ratio as expected from a previous theory [12]. Indeed, for a mixture of crowders at given ϕ_c, the greatest stabilization will occur when all the crowders are of the more stabilizing (smaller) type.

From our explicit dynamical model of folding we are able to estimate folding rates directly from mean first passage time calculations. Starting from at least 400 different initial coordinates drawn from an equilibrium unfolded ensemble ($Q = 0.2$) at a given packing

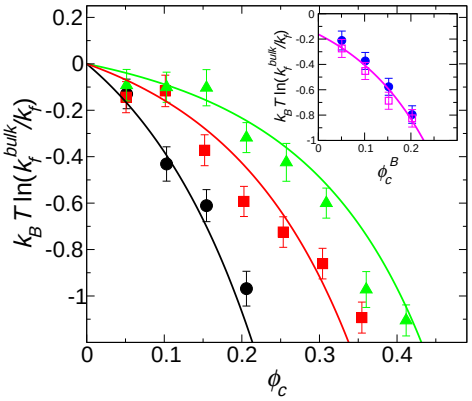

FIG. 4: Influence of crowding on folding kinetics. The effect of crowding on the folding barrier height of prb is estimated using $k_{\mathrm{B}}T\ln(k_{\mathrm{f}}^{\mathrm{bulk}}/k_{\mathrm{f}})$ for crowders of radius 8 Å (circle), 12 Å (square) and 16 Å (triangle up); lines are the SPT predictions. *Inset*: Folding rates with binary crowder mixtures (A: 8 Å radius, B: 12 Å radius, $\phi_c^A = 0.05$). Results from direct simulations (circle) are compared with the predictions based on simulations with a single crowder size (square) and from SPT (line).

fraction, we calculate the average time $\tau_{\mathrm{f}} = 1/k_{\mathrm{f}}$ taken to reach the folded state ($Q > 0.9$). We estimate the change in barrier height ($\Delta F_{\ddagger-\mathrm{U}} = F_{\ddagger} - F_{\mathrm{U}}$; $F_{\ddagger} = F(Q_{\ddagger})$) with crowding as $k_{\mathrm{B}}T\ln(k_{\mathrm{f}}^{\mathrm{bulk}}/k_{\mathrm{f}})$, which is justified if the position of the folding barrier and the diffusion coefficient along the reaction coordinate are unchanged (Figure 4). The folding barrier height decreases monotonically with increasing ϕ_c: indeed, it is possible to predict the change in rates using SPT with no further adjustable parameters: we calculate the transition state radius $a_{\ddagger}$ directly from the Q umbrella simulations ($Q = 0.55$), and estimate the change in barrier height from SPT, using $a_{\ddagger}$ in place of a_{N}. The effect of mixed macromolecular crowding on kinetics can also be obtained from single crowder simulations by assuming additivity (Figure 4, inset).

In summary, we find that scaled particle theory provides an accurate description of the effect of macromolecular crowding on both folding stability and rates. The effect of purely repulsive crowders can be well-approximated over a wide range of crowder sizes and packing fractions by treating the unfolded state as a hard sphere of fixed radius a_{U}. In all the cases considered, we find $a_{\mathrm{U}} \approx 0.8R_{\mathrm{g}}^{\mathrm{U}}$. Our results have several important consequences. For crowders of a single size, folding rate and stability will increase with increasing packing fraction monotonically under the relevant physiological conditions. When considering mix-

tures of purely repulsive crowders of different sizes the effects of crowding are additive, and the most stabilizing composition will consist completely of the smallest crowder. Therefore any maximum in stabilization as a function of the ratio of various components at a fixed total packing fraction implies the existence of attractive interactions with at least one of the crowders [32].

We are grateful to Dr. Attila Szabo for several helpful discussions. This research was supported by the Intramural Research Program of the NIH, NIDDK. J.M. thanks Dr. Artur Adib for supporting a postdoctoral fellowship. R.B. is supported by a Royal Society University Research Fellowship. We gratefully acknowledge the Human Frontier Science Program short-term fellowship to J.M. for which this work was originally proposed. This study utilized the high-performance computational capabilities of the Biowulf PC / Linux cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov), Theory Sector computing resources at Cambridge, and the NSF TeraGrid resources provided by TACC.

---

* jeetain@helix.nih.gov

† rbb24@cam.ac.uk

[1] R. J. Ellis and A. P. Minton, Nature **425**, 27 (2003).

[2] A. P. Minton, Mol Cell. Biochem. **55**, 119 (1983).

[3] M. S. Cheung, D. Klimov, and D. Thirumalai, Proc. Natl. Acad. Sci. USA **102**, 4753 (2005).

[4] M. S. Cheung and D. Thirumalai, J. Phys. Chem. B **111**, 8250 (2007).

[5] D. D. Minh, C. E. Chang, J. Trylska, V. Tozzini, and J. A. McCammon, J. Am. Chem. Soc. **128**, 6006 (2006).

[6] L. Stagg, S. Q. Zhang, M. S. Cheung, and P. Wittung-Stafshede, Proc. Natl. Acad. Sci. USA **104**, 18976 (2007).

[7] H.-X. Zhou, G. Rivas, and A. P. Minton, Ann. Rev. Biophys. **37**, 375 (2008).

[8] J. K. Cheung, V. K. Shen, J. R. Errington, and T. M. Truskett, in *Statistical Mechanics of Cellular Systems and Processes*, edited by M. H. Zaman (Cambridge University Press, Cambridge, 2009).

[9] J. L. Lebowitz and J. S. Rowlinson, J. Chem. Phys. **41**, 133 (1964).

[10] A. P. Minton, Biophys. J. **88**, 971 (2005).

[11] H.-X. Zhou, Arch. Biochem. Biophys. **469**, 76 (2008).

[12] H.-X. Zhou, Proteins **72**, 1109 (2008).

[13] T. Wang, Y. Zhu, and F. Gai, J. Phys. Chem. B **108**, 3694 (2004).

[14] S. J. Hamill, A. Steward, and J. Clarke, J. Mol. Biol. **297**, 165 (2000).

[15] E. L. McCallister, E. Alm, and D. Baker, Nat. Struct. Biol. **7**, 669 (2000).

[16] H. Nymeyer, A. E. García, and J. N. Onuchic, Proc. Natl. Acad. Sci. U. S. A. **95**, 5921 (1998).

[17] M. Mickler, R. I. Dima, H. Dietz, C. Hyeon, D. Thirumalai, and M. Rief, Proc. Natl. Acad. Sci. U. S. A. **104**, 2026820273 (2007).

[18] J. Karanicolas and C. L. Brooks, III, Prot. Sci. **11**, 2351 (2002).

[19] P. G. Wolynes, Q. Rev. Biophys. **38**, 405 (2005).

[20] J.-E. Shea, J. N. Onuchic, and C. L. Brooks, III, J. Chem. Phys. **113**, 7663 (2000).

[21] Y. Levy, P. G. Wolynes, and J. N. Onuchic, Proc. Natl. Acad. Sci. U.S.A. **101**, 511 (2004).

[22] L. L. Chavez, J. N. Onuchic, and C. Clementi, J. Am. Chem. Soc. **126**, 8426 (2004).

[23] R. B. Best and G. Hummer, Proc. Natl. Acad. Sci. U. S. A. **102**, 6732 (2005).

[24] C. L. Brooks, III, A. Brünger, and M. Karplus, Chem. Phys. Lett. **105**, 495 (1984).

[25] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comp. Chem. **4**, 187 (1983).

[26] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).

[27] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, J. Comp. Chem. **13**, 1011 (1992).

[28] R. B. Best and G. Hummer, Phys. Rev. Lett. **96**, 228104 (2006).

[29] J. Mittal and R. B. Best, Proc. Natl. Acad. Sci. U. S. A. **105**, 20233 (2008).

[30] S. Asakura and F. Oosawa, J. Chem. Phys. **22**, 1255 (1954).

[31] A. A. Louis, P. G. Bolhuis, J. P. Hansen, , and E. J. Meijer, Phys. Rev. Lett. **85**, 2522 (2000).

[32] F. Du, Z. Zhou, Z.-Y. Mo, J.-Z. Shi, J. Chen, and Y. Liang, J. Mol. Biol. **364**, 469 (2006).